DOCUMENT RESUME

ED 395 992 TM 025 186

AUTHOR Wang, Xiang Bo
TITLE Understanding Psychological Processes That Underlie Examinees' Choices of Constructed Response Items on a Test.
PUB DATE 9 Apr 96
NOTE 33p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (New York, NY, April 9-11, 1996).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150) -- Tests/Evaluation Instruments (160)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Chemistry; *College Students; *Constructed Response; Curriculum; Essay Tests; Higher Education; High Schools; *High School Students; Performance Based Assessment; *Psychological Patterns; *Student Attitudes; *Test Construction; Test Items
IDENTIFIERS Advanced Placement Examinations (CEEB); *Choice Behavior

ABSTRACT

It has been found that when examinees are allowed to choose a subset of constructed response (CR) items to answer on a test, they tend to choose differently and often perform lower on more popularly chosen items. The purpose of this study was to examine this finding. Using an experiment that incorporated a revised Advanced Placement (AP) Chemistry examination and a well-designed questionnaire, this study revealed a series of psychological factors that consistently influence examinee's choices of CR items. The instrument package was administered to 680 students in Hawaii at 4 levels of college and high school chemistry study. The national choice pattern of the five essays was well replicated by the Hawaiian students, indicating the existence of a consistent influence on examinees' choices. This influence was attributed to the commonality of chemistry curricula and textbooks. Students' perception of essay difficulty did accurately predict and account for the popularity of essay choices, and their perception of content dimensionality of essays coincided with their choices. The findings from this study offer a number of suggestions regarding CR item pretesting and test construction and a wide range of application possibilities for performance-oriented examinations. An appendix contains the survey and test kit. (Contains 3 tables, 4 figures, and 15 references.) (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it
☐ Minor changes have been made to improve reproduction quality
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
XIANG BO WANG
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

ED 395 992

UNDERSTANDING PSYCHOLOGICAL PROCESSES THAT UNDERLIE EXAMINEES' CHOICES OF CONSTRUCTED RESPONSE ITEMS ON A TEST

Xiang Bo Wang
Psychometrics
Law School Admission Council

ABSTRACT

It has been found repeatedly that when examinees are allowed to choose a subset of constructed response (CR) items to answer on a test, they tend to choose differently and often perform lower on more popularly chosen items. The purpose of this study is to examine this finding. Using an experiment that incorporates a revised Advanced Placement Chemistry Exam and a well-designed questionnaire, this study has revealed a series of psychological factors that consistently influence examinees' choices of CR items. The findings from this study offer a number of suggestions regarding CR item pre-testing and test construction and a wide range of application possibilities for performance-oriented exams.

TM025186

BEST COPY AVAILABLE

ACKNOWLEDGEMENTS

The author wishes to thank:

(1) Dr. Peter Dunn-Rankin, Dr. Harold Ayabe, Dr. Ronald Heck, Dr. Frank Pottenger, and Dr. Shuqiang Zhang of University of Hawaii for their long-term encouragement, guidance and comments;

(2) Dr. Howard Wainer of the Educational Testing Service, and Dr. David Thissen of the University of North Carolina for their valuable advice on item response theory issues;

(3) Mr. Walter MacDonald, Dr. Rick Morgan and Mrs. Behroz Manecshana of the College Board for the releaseve of the 1989 AP Chemistry Exam data;

(4) Dr. Gordon Barrow, Dr. John Gilje, Dr. Randy Larsen, Dr. McDonald Ray of the Chemistry Department, University of Hawai'i for data collection on UH students;

(5) Mr. Mel Ishii of Aiea High School; Mrs. Pam Fujinaka, Iolani School; Mrs. Maria Mendoza and Mrs. Amy Watanabe from Kailua High School; Mrs. Vickie Kirahira of Kaiser High School; Mr. Dennis Otta, Kalani High School; Mrs. Jene Kullingham and Mr Jerry Johnson of Kamehameha High School; Mrs. Katherine Taniguchi of McKinley High School; Mr. Edward Ginoza from Maui High School; Mrs. Ruth Fukuchi of Mililani High School; Mrs. Alice Kaneko, Moanalua High School; Mr. Ted Brattstrom of Pearl City High School; Mr. Michael Hu from Punahou School; Mrs. Sharlene Higa of Roosevelt High School; Mrs. Nancy Rocheleau of the Sacret Hearts Academy for their whole-hearted support in data collection on more than 600 high school students in Hawai'i.

UNDERSTANDING PSYCHOLOGICAL PROCESSES THAT UNDERLIE EXAMINEES' CHOICES OF CONSTRUCTED RESPONSE ITEMS ON A TEST

Xiang Bo Wang
Psychometrics
Law School Admission Council

INTRODUCTION

Although it is increasingly popular for many performance oriented tests to contain constructed response (CR) items, a subset of which can be chosen by examinees to answer, it is only recently that some disturbing facts about the consequences of such choices on test performance have come to light. Substantial numbers of examinees perform poorly on the items they choose (Fremer, Jackson & McPeek, 1968; Pomplum, Morgan & Nellikunnel, 1992). Examinees of different gender and ethnic background seem to choose differently, which results in score biases to their disadvantage (Wainer & Thissen, 1993; Wainer & Thissen, 1994). In order to alleviate such biases, it has been advocated that scores of differentially chosen CR items be compared and equated (Wainer, Wang & Thissen, 1994). The equating theory, methodology and some intestable assumptions have been investigated and explicated (See Wainer & Thissen, 1993; Wainer & Thissen, 1994; Thissen, Wainer & Wang, 1994; Wang, Wainer & Thissen, 1995).

Based on systematic analyses of Part D of the 1989 Advanced Placement (AP) Chemistry Exam this author (Wang, 1996a) reported five surprising findings: (1) The five essays in Part D were chosen in dramatically different ways; (2) The more frequently chosen essays belonged to the core chemistry content, while the least frequently chosen item addressed a highly specialized chemistry topic; (3) Examinees tended to score lower on the more popularly chosen core-chemistry items than on the non-core chemistry items; (4) The order in which the essays were presented seemed to have a significant effect on examinee choice patterns — the examinees who chose items selectively performed significantly better than did those that chose items sequentially; (5) Except for extremely low-ability examinees, all examinees seemed to choose in a similar way.

Although curricular and instructional explanations were offered to the above seemingly contradictory findings through a survey of AP chemistry teachers in the state of Hawaii, there are still several unanswered questions: Why did the national examinees choose the five essays so differently? Can the diverse choice patterns of the national AP chemistry population be replicated by some other independent sample? What role did examinees' perceptions of the difficulty and dimensionality of these essays play in their choices? As a sequel to the earlier study (Wang, 1996a), the purpose of this research is to find answers to these questions with emphasis on the psychological processes underlying examinees' choices.

It should be pointed out that the author makes reference to the findings of the earlier study (Wang, 1996a) which will not be duplicated here for the sake of brevity. The reader is strongly encouraged to read this earlier paper, should any questions arise.

RESEARCH INSTRUMENT, SUBJECTS AND METHODOLOGY

In order to fulfill the above research purpose, this study employed a special experiment that consisted of four parts of questionnaires and tests. As shown in Appendix I, the main instrument used in this experiment was called "Advanced Placement Chemistry Survey and Test Kit" (the Kit hereafter). The Kit consisted of four parts. Parts 1, 2 and 4 are described, as Part 3 was designed to address a different research question which will be reported in a different paper. Part 1 was a general information survey that sought some demographic information on participants, such as age, gender, interest and length of chemistry study, career choices, and so on. Table 1 summarizes the ethnic and gender composition of the 680 Hawaii students participating in this study.

Designed to compare the chemistry ability of the Hawaii participants with the 1989 national AP chemistry exam examinees, Part 2 was a mini AP chemistry test whose twelve MC items were carefully selected from Part A of the 1989 AP Chemistry Exam to mimic its distributions of item content, difficulty and discrimination, as well as its test information. Note the reason that only 12 MC items were used was due to the limited time available to the Hawaii students. These 12 MC items, along with 8 MC items in Part 3, appeared satisfactory for the purpose of this study.

Entitled "AP Chemistry Multiple Choice (MC) Item Comparison and Performance," Part 3 was composed of four pairs of MC items of varying item difficulties, content and discrimination accompanied by a series of questions. The direct relevance of Part 3 to this paper is that its 8 MC items were used along with the 12 items in Part 2 to compare the ability distributions of Hawaii students and national examinees.

Part 4, "AP Chemistry Essay Problem Comparisons" consisted of the same five essays that constituted Part D of the 1989 AP Chemistry Exam. A set of comparison questions was presented in this part. The purpose of Part 4 was to verify whether or not Hawaii participants would independently replicate the national choice patterns of the five CR items found in the 1989 AP chemistry data. A positive finding would strengthen the hypothesis that a systematic influence did underlie the seemingly divergent choices of the 1989 national AP chemistry examinees.

With the support of a wide range of chemistry teachers and students in Hawaii, the Kit was administered to 680 students in Hawaii at four academic levels of chemistry study: (1) one upper-division class of 15 students in the Department of Chemistry, University of Hawaii (UH), who had completed at least four semesters of college chemistry; (2) two lower-division chemistry classes of 81 students in the Chemistry Department, who had completed their first year of college chemistry; (3) thirteen high school AP Chemistry classes of 237 students, a majority of the registered AP Chemistry population throughout the state of Hawaii in 1992; and (4) eleven high school general chemistry classes of 334 students.

Although participation was voluntary, the Hawaii high school chemistry students seemed to perform on more carefully on the Kit than did the UH chemistry students, probably because AP chemistry was more relevant to the former than to the latter. A certain number of students, mostly the UH lower division chemistry students, did not completely finish the survey Kit due to lack of sufficient time, or from academic pressures such as examinations in other courses. Moreover, a small number of students supplied uniform answers on their answer sheets, such as bubbling all "B" options, and so on. To preclude the undue influence of incomplete or random data, those who did not complete Part C of the Kit, or those who supplied uniform answers, were elimininated from this study. As a result, sixty

subjects were deleted from the total sample pool. All the analyses on Hawaii students reported in the remainder of the dissertation are based on 618 cases. Table 4.5 summarizes the main demographic information of the 618 Hawaii students.

Analysis methodology employed both uni- and multi-dimensional scaling methods to the comparison data to reveal participants' perceptions of item difficulty, similarity and dimensionality. Item response theory was used for the calibration and construction of the mini-AP chemistry test.

ANALYSES AND RESULTS

The analyses of this paper is divided into four components. Component I describes the ability comparability between the Hawaii participants and the national AP chemistry population. Component II compares the similarities of the choices between the Hawaii and national groups. Component III assessed students' perception of essay difficulty and the effect of difficulty on essay choices. Component IV investigates on the dimensionality of the five essays, and its relationship to the choices.

Component I: Comparing the Chemistry Ability between the Hawaii Participants and National AP Chemistry Population:

Twenty MC items were selected from the original 75 MC items of Section I of the 1989 AP Chemistry Examination and built into Parts 3 and 4 of the Kit. Altogether, 618 students responded to the 20 items. Table 2 summarizes their performance on the 20 items in reference to the national norm.

Insert Table 2 here.

The AP chemistry students in Hawaii scored about the same as the national norm. As expected, since most of the non-AP students had completed only about one year of general chemistry study by the time they responded to this instrument, their mean performance on this instrument was only half that of their AP counterparts. The mean score for the UH upper-division class students was 12.36. The mean of the UH lower division chemistry students was 5.73.

In summary, the overall performance of Hawaii students on the 20 MC items was quite similar to that of the national norm. These 20 MC items functioned equally consistently to both Hawaii and national students as shown by 0.76 KR-20 coefficients for both groups. Figure 1 illustrates detailed score distributions for the four major groups of students.

Insert Figure 1 here.

Component II: Comparing Essay Choices between Hawaii Students and National AP Chemistry examinees:

Given that the Hawaii students performed so closely to the national norm on average, would they choose the five essays in a similar way as did the 1989 AP Chemistry examinees? Note that the students in Hawaii were asked only to indicate in Part D of the Kit how they would like to choose three of the five essays. Figure 2 reveals strikingly familiar essay choice patterns between the Hawaii students and the national examinees. Based on 554 Hawaii students, the overall choice pattern of the five essays

substantially mirrored that of the 18,462 examinees who took the 1989 AP Chemistry Exam. The correlation between the two patterns is .87.

Insert Figure 2 here.

Three points should be noted. First, like the 1989 AP Chemistry Exam population, the Hawaii students also liked Essay Combination 5, 6 & 8 best. Essay Combination 5, 7 & 8 remained the second most popular combination. Second, almost the same proportion of Hawaii students (13%) as the 1989 AP examinees (14%) chose Essay Combination 5, 6 & 7. Note that Essays 7, originally the third essay in the 1989 AP Chemistry Exam, was presented as the fifth essay in the Kit to test the effect of positioning Essay 9. In spite of the position change, Essay 7 was still preferred over Essay 9. There seems to be something inherent about Essay Combination 5, 6 & 7 that attracts AP chemistry students.

The above findings infer a universal regularity in the way students choose these essays, possibly due to the commonality of the AP chemistry textbooks. Although there is no standardized national AP Chemistry curriculum, all AP chemistry textbooks are quite similar in their curricular emphases, which transcends into students' varying familiarity with various subjects, and eventually the way they choose items on a test.

Component III: Relationship between Students' Perceptions of Essay Difficulty and Their Choices

In light of the high levels of similarity in choice tendencies involving the five essays, one might wish to discern the cognitive dimensionality underlying their choices. The first questions are: "How did the students perceive the difficulties of the five essays? How did their perception influence their choices?"

In Part D of the Kit, Hawaii students were asked to conduct pair-wise comparisons of the relative difficulty levels of the five essays through the question "Which one (essay) seems easier for you?" Such pair-wise comparison data were analyzed through Ranko (Dunn-Rankin, 1983), which carries out variance stable rank scaling analysis. The linear plot with scale scores from RANKO is reproduced in Figure 3.

Insert Figure 3 here.

Hawaii students ranked Essay 5 as the easiest, followed by Essays 8, 6, 7 and 9 in that order. Such an order of perceived essay difficulty conforms completely with the popularity of the five essays of the 1989 national chemistry examinees reported in Wang's earlier study (See Figure 2 of Wang, 1996a). Using this rank order, one also can predict the popularity of the ten essay combinations found with the national examinee population (see Figure 1 of Wang, 1996a). For example, from Figure 2 of this study, we know that Essays 5, 6 & 8 form the most popular essay combination, which happened to be the three easiest essays on the Ranko scale. From the Ranko scale, we can also predict that Essays 5, 8 & 9 and 5, 6 & 7 form the second and third most popular essay combinations. The three essays that form the least popular Essay Combination 6, 7 & 9 in Figure 2 turn out to be the last three essays in the Ranko scale. Table 3 further summarizes the critical differences among the essays, and all the essays are shown to be significantly different from one another in terms of their difficulty at 0.01 significance level.

Insert Table 3 here.

Component IV: Relationship between Students' Perceptions of Content Dimensionality and Their Choices

These five essay questions are known to involve the following areas of chemistry knowledge:

Essay 5:	valence, electronic configuration, covalent bonding, molecular geometry
Essay 6:	periodic trends, stability, ionization, energy, properties of halogens, properties of alkali metals
Essay 7:	properties of metals, writing and balancing chemical equations, conservation of mass, double displacement reactions
Essay 8:	rates of reaction, physical behavior of gases, energy changes
Essay 9:	nuclear chemistry

It is clear that these five essay questions differ in content. How did such content diversity affect examinees' choices?

Because not all examinees responded to these five essays, traditional factor analysis based on students' scores cannot be applied here to ascertain the dimensionality of the five essays. In order to offset this information deficit, pair-wise similarity comparisons ("How similar are Problem I and II?") have been incorporated into Part D of the AP Chemistry Survey and Test Kit. Based on the similarity data from the Hawaii students, the dimensionality in terms of content similarity of the five essays is revealed through multidimensional scaling in Figure 4.

Insert Figure 4 here.

To correspond to content differences, these five essays are spread out in the four quadrants of the two dimensional space. However, under the seemingly large differences lies a certain vein of commonality. First, if we examine the two halves divided by the vertical line, we see that Essays 5, 6 & 8 are in the left half, and Essays 7 and 9, in the right half. What do Essays 5, 6 & 8 share in common? They reflect the most frequently taught topics and constitute the common core of general chemistry. On the other hand, Essay 7 requires more extensive lab experience and not just classroom lecturing. Since not all students have equal access to laboratories, Essay 7 cannot be dealt with as readily by average students. Furthermore, since Essay 9 is concerned with the least taught topic of nuclear chemistry, it is the further removed from the other four essays.

Figure 4 designates the horizontal dimension as "extent of textbook coverage" with the left half symbolizing "core chemistry" topics and the right half, the "non-core chemistry" topic.

How can the vertical dimension be interpreted? The Hawaii AP chemistry teachers survey (Wang, 1996a) indicates that both Essays 6 and 7 tap into deeper and more complex chemistry theories, structures and lab experiences, while Essays 5, 8 & 9 are relatively descriptive and fact-oriented. Therefore, The vertical dimension denotes "complexity of problems" with the upper half standing for the relatively "straightforward" questions, while the lower half represents relatively "complicated" questions.

How did the dimensionality affect examinees' choices? More specifically, did students tend to choose essays of the same or similar dimension? The answer appears to be *Yes*. For example, according to

Figure 2, Essay Combination 5, 6 & 8 was the most frequently chosen combination by both the national population and the Hawaii students. These three essays form the "core chemistry" half of the horizontal dimension. Any essay combination with Essays 7 & 9 was usually avoided. Essays 6, 7 & 9 were the least favored combination, basically because Essays 7 & 9 form the "non-core chemistry" half of the horizontal dimension.

Moreover, as shown in Table 4 of the previous study (Wang, 1996a), we know that Essay Combination 5, 8 & 9 has the highest mean score. This is probably because these three essays are from the "straightforward" half of the vertical dimension. It also can be seen that any essay combination that includes Essays 6 & 7 has a lower mean score. For instance, Essay Combination 5, 6 & 7 has the lowest mean score of the ten essay combinations, basically because Essays 6 & 7 were from the "complicated" dimension of the essays.

It can be concluded from the above findings that not only were students' choices influenced by item dimensions, but also their scores. The origin of essay dimensions can be attributed to various factors. In the case of these five essays, it is reasonable to believe that these two dimensions are attributed to the order of textbook presentation and tasks involved to solve the problems. It is argued here that with most subject tests like AP chemistry tests, the predominant mode of textbook presentation must have had a long-lasting effect on how students are taught, which subsequently influences how they will choose on a test.

CONCLUSION AND DISCUSSION

This study has produced three major findings. First, the national choice pattern of the five essays was well replicated by the participants of Hawaii, indicating the existence of a consistent influence on examinees' choices. This consistent influence is attributed to the commonality of AP chemistry curriculums and textbooks. The second finding is that students' perception of essay difficulty, when transformed into a rank order, can accurately predict and account for the popularity of essay choices. The third finding is that students' perception of content dimensionality of the essays coincides with students' choices, and explains their mean performance on the essay combinations. The second and third findings vividly suggest the links between students' psychological processes and their choices.

The above findings offer at least two useful avenues to explore how to better implement CR item choices on a test. The first useful application is to pre-assess how likely examinees would choose a set of CR items on a test. It is well-known to testing agencies that most CR items are difficult to pre-test because of logistical difficulties in scoring them and the high risk of test security breaches. Yet, this paper shows that having a small number of potential examinees to rate difficulty and similarity levels can offer fairly accurate estimates of how examinees perceive them in terms of their difficulties and dimensionality, and of how likely a candidate would choose them. This method would minimize pretesting costs and test security risks.

The second possible application is to minimize the potential differences among CR items. The principle of test equity demands a fair and equal chance of success for each test taker. Yet, allowing examinees to choose among a set of CR items potentially different in difficulty, content and dimensionality would easily jeopardize such a principle. Controlling content similarities is certainly one solution to this equity.

However, it is well known that content-similar items frequently produce psychologically different dimensions. The technology used in this study can certainly help reveal such psychological dimensionality differences to further improve the quality of the test.

Note that it has been found in this paper that students' perceptions of item difficulty and dimensionality can account for their choices, it does not mean that such perceptions can accurately predict the actual difficulty of a test item. According to Wang (1996a), examinees on the national level performed lower on the more frequently chosen items. Why was there a negative relationship between familiar items and performance? The investigation of this fallacy will be reported in a separate study. It suffices to say that there seemed to exist a negative interaction between the scoring rubrics and the curricular emphases on the essay items. More specifically, more commonly taught chemistry items might have been scored more stringently than less commonly taught essay items. Through an experiment with MC items not requiring scoring rubrics but straightforward "right" or "wrong" scoring, examinees do score higher on the items they choose, as long as they understand them.

This study does have its shortcomings. The author would have preferred a more standard and controlled fashion of delivery for the Kit. More reliable and complete responses would have been obtained, if students responded to the Kit under some mandatory condition. However, best efforts were made to collect the data, given all the resources available to the author, and the results have been instructive. It is hoped that this study will stimulate more research to help to increase the accuracy, reliability and validity of tests that involve CR item choices.

REFERENCES

The College Board. (1990). The 1989 Advanced Placement Examination in Chemistry and Its Grading. New York: Author.

Dunn-Rankin, P. (1983). Scaling methods. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.

Fremer, J., Jackson, R., & McPeek, M. (1968). Review of the psychometric characteristics of the Advanced Placement tests in Chemistry, American History, and French (Unpublished Tech. Memorandum). Princeton, NJ: Educational Testing Service.

Keppel, G.(1982). Design & Analysis, a Researcher's Handbook. Prentice-Hall, Inc., Englewood Cliffs, New Jersey 07632.

Masterton, W. L., Slowinski, E. J., & Stanitski, C. L. (1985). Chemical Principles, 6th edition. Saunders College Publishing, Philadelphia.

Pomplum, M., Morgan, R., & Nellikunnel, A. (1992). Choice in advanced placement tests (unpublished statistical Rep. No. SR-92-51). Princeton, NJ: Educational Testing Service.

Thissen, D., Wainer, H., & Wang, X-B. (1994). Are tests comprising both multiple choice and free-response items necessarily less unidimensional than multiple-choice tests? Analysis of two tests. *Journal of Educational Measurement, 31,* 113-123.

Wainer, H., & Thissen, D. (1992). Choosing A Test. (Tech. Rep. No. 92-25). Princeton, NJ: Educational Testing Service.

Wainer, H., & Thissen, D. (1993). Combining multiple choice and constructed response test scores: toward a Marxist theory of test construction. *Applied Measurement in Education, 6,* 103-118.

Wainer, H., & Thissen, D. (1994). One examinee choice in educational testing. *Review of Educational Research, 64,* 159-195.

Wainer, H., Wang, X-B., & Thissen, D. (1994). How well can we compare scores on test forms that are constructed by examinees' choice? *Journal of Educational Measurement, 31,* 183-199.

Wang, X. B., Wainer, H., & Thissen, D. (1995). On the viability of some untestable assumptions in equating exams that allow examinee choice. *Applied Measurement in Education, 8(3),* 211-225.

Wang, X. B. (1996a). Investigating allowing examinees to choose constructive response items on a test. Paper accepted by 1996 AERA Annual Meeting, New York, April 8-12.

BEST COPY AVAILABLE

Wang, X. B. (1996b). Understanding psychological processes that underlie examinees' choices of constructed response items on a test. Paper accepted by 1996 NCME Annual Meeting, New York, April 8-12.

Wittrock, M., & Baker, E., (1992). *Testing and Cognition*. Prentice Hall, Englewood Cliffs, NJ 07632.

Table 1: Summary of Demographic Information on Survey Subjects

	Category	Frequency	Percent	Cumulative Frequency	Cumulative Percent
GENDER	Male	262	42.4	262	42.4
	Female	314	50.8	576	93.2
	Unidentified	42	6.8	618	100.0
ETHNICITY	Japanese	211	34.1	211	34.1
	Chinese	139	22.5	350	56.6
	Caucasian	75	12.1	425	68.7
	Filipino	74	12.0	499	80.7
	Hawaiian or Part-Hawaiian	38	6.2	537	86.9
	Portuguese	6	1.0	543	87.9
	Black	3	0.5	546	88.4
	Samoan	1	0.2	547	88.6
	Other Asians	47	7.6	594	96.2
	Others	21	3.4	615	99.6
	Identified	3	0.4	618	100.0
STUDENT LEVEL	AP Chemistry	237	38.3	237	38.3
	College	48	7.8	285	46.1
	Non-AP	333	53.9	618	100.0
Total Hawaii Student #		618			

Note: AP stands for Advanced Placement.

Table 2: Comparing AP Chemistry Performance

between Hawaiian Students and 1989 AP Chemistry Examinees

Student Category	Student #	Min. Score	Max. Score	Mean	StD	KR-20	SEM
HI AP Chemistry	237	1	20	8.72	4.41	.83	2.04
Non-AP Chemistry	333	0	16	4.96	2.11	.26	1.92
UH Upper Division	15	6	20	12.36	4.09	.83	2.04
UH Lower Division	33	2	15	5.73	1.98	.52	1.98
Total Hawaii Sample	618	0	20	7.94	3.78	.76	1.97
National AP Sample	1000	0	20	8.42	3.66	.76	2.10

Table 3: Rank Differences

Essays	5	8	6	7	9
5	0				
8	565	0			
6	735	170	0		
7	927	362	192	0	
9	1108	543	373	181	0

Note: The critical differences are 137 at .05 level and 163 at .01 level.

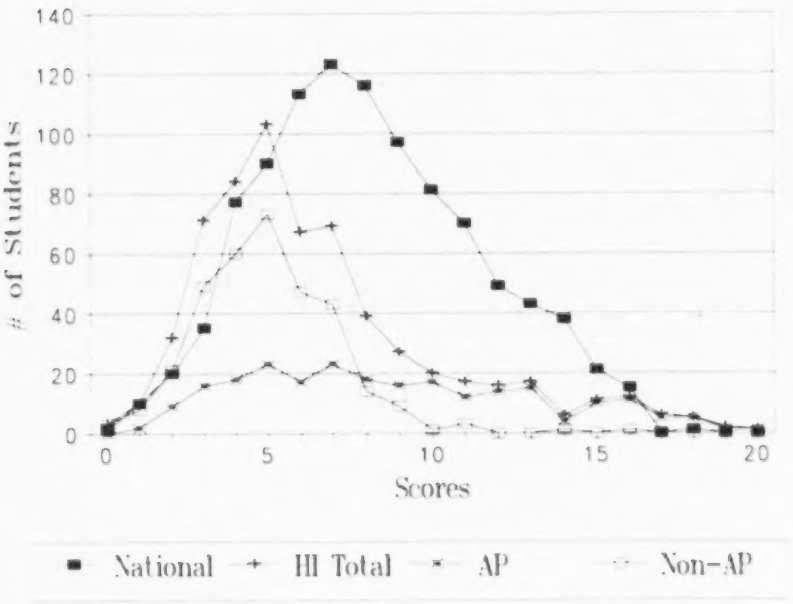

Figure 1: Raw Score distributions of National, Hawaii Total, AP and Non-AP Samples

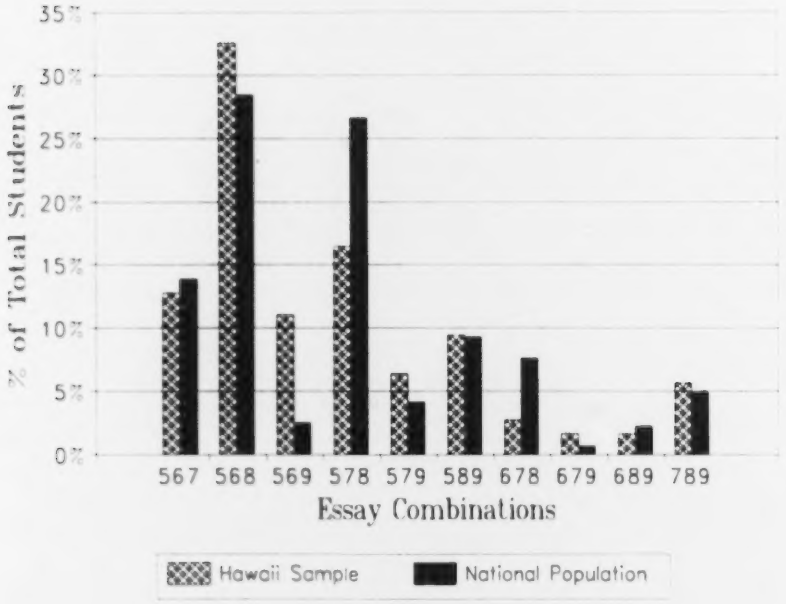

Figure 2: Comparison of the Overall Essay Choice Pattern of 554 Hawaii and 18,462 National Students

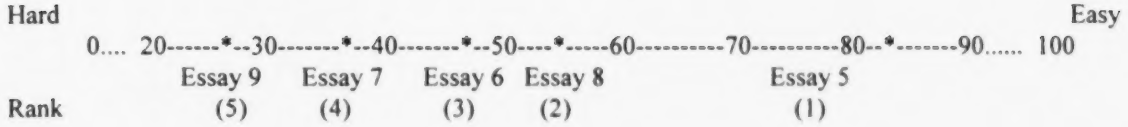

Figure 3: Perceived Order of Essay Difficulty

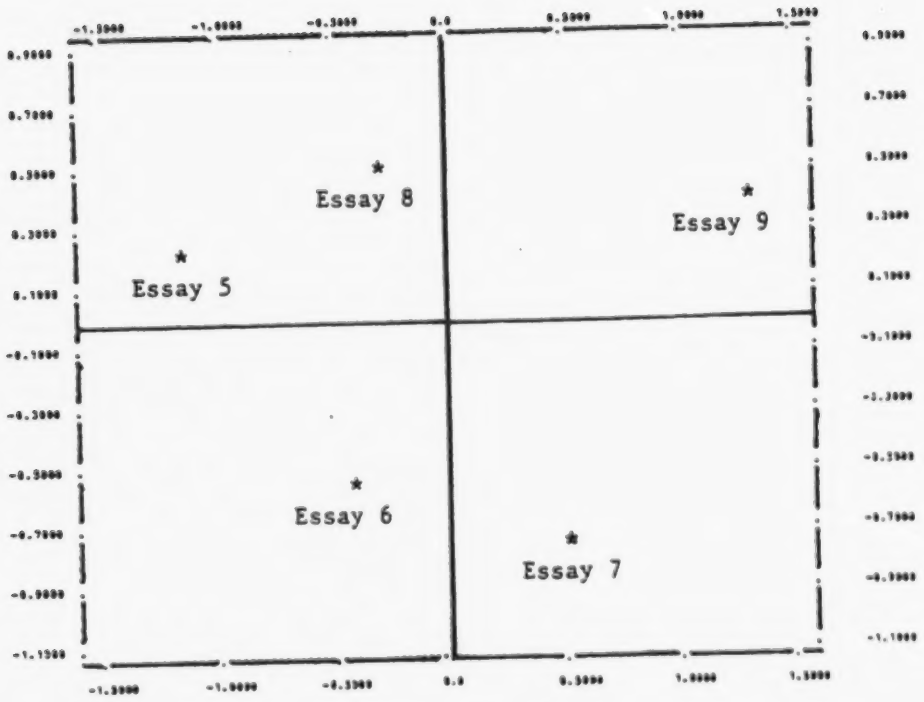

Figure 4: Dimensionality of the Five Essay Questions

BEST COPY AVAILABLE

APPENDIX 1

ADVANCED PLACEMENT

CHEMISTRY

SURVEY AND TEST KIT

ADVANCED PLACEMENT

CHEMISTRY

SURVEY AND TEST KIT

General Instructions: This instrument contains *BOTH* survey and test questions. It is *NOT* an ordinary multiple choice test. It is *IMPORTANT* that you read the instructions for each section carefully in order to provide accurate answers. All the questions are numbered continuously. Please indicate your answer by completely filling in the corresponding oval on the supplied answer sheet. Give only one answer to each question.

Please note: Advanced Placement Chemistry is abbreviated as AP Chemistry throughout this kit.

Please do *NOT* make marks in this kit!

Please *TURN IN* this kit with your answer sheet!

THANK YOU VERY MUCH FOR YOUR COOPERATION!

Part A: General Information

Instructions: Please take out the answer sheet supplied to you and answer the following questions to the best of your knowledge. You may ignore the optional questions. Please note that the answers to the first five questions go on the *left half* of the answer sheet.

I. Your name (Optional)

II. Sex

III. Grade (Note: For college students, Freshman = 13 Sophomore = 14 Junior = 15 Senior = 16)

IV. Birthday (Optional)

V. ID Number (SS#)

VI. Special Code: With which of the following groups do you most closely identify?

(A) Caucasian
(B) Japanese
(C) Filipino
(D) Hawaiian & Part-Hawaiian
(E) Chinese
(F) Samoan
(G) Portuguese/Spanish
(H) Black
(I) Other Asian
(J) Others

1. Have you ever taken an AP Chemistry course?

(A) ***Yes*** (B) ***No***

2. Do you consider an AP Chemistry course useful?
(A) ***Yes*** (B) ***No***

3. Have you taken an AP Chemistry test?

(A) ***Yes*** (B) ***No***

4. If the answer to Question 3 is *Yes*, do you think it is a good test to reflect your chemistry ability?

(A) ***Yes*** (B) ***No***

Otherwise, mark Oval (E).

5. Do you plan to major in Chemistry in College?

(A) ***Yes*** (B) ***No***

6. Do you plan to obtain a Bachelor of Science degree in College?

(A) ***Yes*** (B) ***No***

7. Do you consider chemistry one of your favorite academic subjects?

(A) ***Yes*** (B) ***No***

8. Do you consider chemistry one of your strongest academic subjects?

(A) ***Yes*** (B) ***No***

9. How many high school chemistry courses have you taken?

(A) ***1*** (B) ***2***
(C) ***3*** (D) ***4***
(E) ***5 or more***

10. How many college chemistry courses have you taken?

(A) ***0*** (B) ***1***
(C) ***2*** (D) ***3***
(E) ***4*** (F) ***5***

Part B: Mini AP Chemistry Test

Instructions: **For the following 11 multiple choice problems (11-21), please choose the most appropriate answer. For all problems involving solutions and /or chemical equations, assume that the system is in water and at room temperature unless otherwise stated. Feel free to use the attached tables whenever necessary.**

11. Which is the most electronegative element?

    (A) O (B) La (C) Rb
    (D) Mg (E) N

12. Which is used to explain the fact that the four bonds in methane are equivalent?

    (A) Hydrogen bonding
    (B) Hybridization

    (C) Ionic bonding
    (D) Resonance
    (E) van der Waals forces (London dispersion forces)

13. $2\ K + 2\ H_2O \rightarrow 2\ K^+ + 2\ OH^- + H_2$

    When 0.400 mole of potassium reacts with excess water at standard temperature and pressure as shown in the equation above, the volume of hydrogen gas produced is

    (A) 1.12 liters (B) 2.24 liters
    (C) 3.36 liters (D) 4.48 liters
    (E) 6.72 liters

14. The $SbCl_5$ molecule has a trigonal bipyramid structure. Therefore, the hybridization of Sb orbitals should be

    (A) sp^2 (B) sp^3 (C) dsp^2

    (D) dsp^3 (E) d^2sp^3

15. A white solid is observed to be insoluble in water, insoluble in excess ammonia solution, and soluble in dilute HCl. Which of the following compounds could the solid be?

    (A) $CaCO_3$ (B) $BaSO_4$
    (C) $Pb(NO_3)_2$ (D) AgCl
    (E) $Zn(OH)_2$

16. The solubility of CuI is 2×10^{-6} -molar. What is the solubility product constant, K_{sp}, for CuI?

    (A) 1.4×10^{-3} (B) 2×10^{-6}
    (C) 4×10^{-12} (D) 2×10^{-12}
    (E) 8×10^{-18}

17. A measured mass of an unreactive metal was dropped into a small graduated cylinder half filled with water. The following measurements were made.

    Mass of metal = 19.611 grams
    Volume of water before addition of metal = 12.4 milliliters
    Volume of water after addition of metal = 14.9 milliliters

    The density of the metal should be reported as

    (A) 7.8444 grams per mL
    (B) 7,844 grams per mL
    (C) 7.84 grams per mL
    (D) 7.8 grams per mL
    (E) 8 grams per mL

18. $H_2PO_4^- + HBO_3^{2-} \rightleftharpoons HPO_4^{2-} + H_2BO_3^-$

The equilibrium constant for the reaction represented by the equation above is greater than 1.0. Which of the following gives the correct relative strengths of the acids and bases in the reaction?

Acids	Bases

(A) $H_2PO_4^- > H_2BO_3^-$ and $HBO_3^{2-} > HPO_4^{2-}$
(B) $H_2PO_3^- > H_2BO_4^-$ and $HBO_3^{2-} > HPO_4^{2-}$
(C) $H_2PO_4^- > H_2BO_3^-$ and $HBO_4^{2-} > HPO_3^{2-}$
(D) $H_2PO_3^- > H_2BO_4^-$ and $HBO_4^{2-} > HPO_3^{2-}$
(E) $H_2PO_4^- = H_2BO_3^-$ and $HBO_4^{2-} = HPO_3^{2-}$

19. $(CH_3)_3CCl(aq) + OH^- \rightarrow (CH_3)_3COH(aq) + Cl^-$

For the reaction represented above, the experimental rate law is given as follows.

$$\text{Rate} = k[(CH_3)_3CCl]$$

If some solid sodium hydroxide is added to a solution that is 0.010-molar in $(CH_3)_3CCl$ and 0.10-molar in NaOH, which of the following is true? (Assume the temperature and volume remain constant.)

(A) Both the reaction rate and k increase.
(B) Both the reaction rate and k decrease.
(C) Both the reaction rate and k remain the same.
(D) The reaction rate increases but k remains the same.
(E) The reaction rate decreases but k remains the same.

20. $MnS(s) + 2H^+ \rightleftharpoons Mn^{2+} + H_2S(g)$

At 25°C the solubility product constant, K_{sp}, for MnS is 5×10^{-15} and the acid dissociation constants K_1 and K_2, for H_2S are 1×10^{-7} and 1×10^{-13}, respectively. What is the equilibrium constant for the reaction represented by the equation above at 25° C ?

21. $H_2O(s) \rightarrow H_2O(l)$

When ice melts at its normal melting point, 273.16k and 1 atmosphere, which of the following is true for the process shown above?

(A) $\Delta H < 0, \Delta S > 0, \Delta V > 0$
(B) $\Delta H < 0, \Delta S < 0, \Delta V > 0$
(C) $\Delta H > 0, \Delta S < 0, \Delta V < 0$
(D) $\Delta H > 0, \Delta S > 0, \Delta V > 0$
(E) $\Delta H > 0, \Delta S > 0, \Delta V < 0$

Part C: AP Chemistry
Item Comparison and Performance

Instructions: **In this section, there are four sets of comparisons of AP Chemistry problems and some questions about them. Please first *read through* the AP Chemistry problems and then answer the questions *in the order they are presented.***

Comparison I

Instructions: Please first examine Comparison Problems I and II below, and then answer questions 22-29 about them.

I. ... $Cr_2O_7^{2-}$ + ... e^- + ... H^+ --> ... Cr^{3+} + ... $H_2O(l)$

When the equation for the half reaction above is balanced with the lowest whole-number coefficients, the coefficient for H_2O is

(A) 2 (B) 4 (C) 6
(D) 7 (E) 14

II. A solution of toluene (molecular weight 92.1) in benzene (molecular weight 78.1) is prepared. The mole fraction of toluene in the solution is 0.100. What is the molality of the solution?

(A) 0.100 *m* (B) 0.703 *m*
(C) 0.921 *m* (D) 1.28 *m*
(E) 1.42 *m*

---

22. To what extent have you studied a problem like Problem I?

(Never) A B C D E (Extensively)

23. To what extent have you studied a problem like Problem II?

(Never) A B C D E (Extensively)

24. How similar are these two questions in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

25. Which of the above two questions seems easier for you?

(A) *Problem I* (B) *Problem II*

26. If you were allowed to choose between these two questions, which one would you choose?

(A) *Problem I* (B) *Problem II*

27. Choose the correct answer to *Problem I*: (A) (B) (C) (D) (E)

28. Choose the correct answer to *Problem II*: (A) (B) (C) (D) (E)

29. If you would like to choose between these two questions *again* after you have tried to answer them, which one would you choose now?

(A) *Problem I* (B) *Problem II*

Comparison II

Instructions: Please first examine Comparison Problems I and II below, and then answer questions 30-37 about them.

I. Which of the following is probably true for a solid solute with a highly endothermic heat of solution when dissolved in water?

(A) The solid has a low lattice energy.
(B) As the solute dissolves, the temperature of the solution increases.
(C) The resulting solution is ideal.
(D) The solid is more soluble at higher temperatures.
(E) The solid has a high energy of hydration.

II. A 2.00-liter sample of nitrogen gas at 27°C and 600. millimeters of mercury is heated until it occupies a volume of 5.00 liters. If the pressure remains unchanged, the final temperature of the gas is

(A) 68°C (B) 120.°C
(C) 477°C (D) 677°C
(E) 950.°C

---

30. To what extent have you studied a problem like Problem I?

(Never) A B C D E (Extensively)

31. To what extent have you studied a problem like Problem II?

(Never) A B C D E (Extensively)

32. How similar are these two questions in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

33. Which of the above two questions seems easier for you?

(A) *Problem I* (B) *Problem II*

34. If you were allowed to choose between these two questions, which one would you choose?

(A) *Problem I* (B) *Problem II*

35. Choose the correct answer to Problem I: (A) (B) (C) (D) (E)

36. Choose the correct answer to *Problem II*: (A) (B) (C) (D) (E)

37. If you would like to choose between these two questions *again* after you have tried to answer them, which one would you choose now?

(A) *Problem I* (B) *Problem II*

Comparison III

Instructions: Please first examine Comparison Problems I and II below, and then answer questions 38-45 about them.

I. Equal volumes of 0.10-molar H_3PO_4 and 0.20-molar KOH are mixed. After equilibrium is established, the type of ion in solution in largest concentration, other than the K^+ ion, is

(A) $H_2PO_4^-$ (B) HPO_4^{2-}
(C) PO_4^{3-} (D) OH^-
(E) H^3O^+

II. Adding water to some chemicals can be dangerous because large amounts of heat are liberated. Which of the following does NOT liberate heat when water is added to it?

(A) KNO_3 (B) NaOH
(C) CaO (D) H_2SO_4
(E) Na

---

38. To what extent have you studied a problem like Problem I?

(Never) A B C D E (Extensively)

39. To what extent have you studied a problem like Problem II?

(Never) A B C D E (Extensively)

40. How similar are these two questions in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

41. Which of the above two questions seems easier for you?

(A) *Problem I* (B) *Problem II*

42. If you were allowed to choose between these two questions, which one would you choose?

(A) *Problem I* (B) *Problem II*

43. Choose the correct answer to *Problem I*: (A) (B) (C) (D) (E)

44. Choose the correct answer to *Problem II*: (A) (B) (C) (D) (E)

45. If you would like to choose between these two questions *again* after you have tried to answer them, which one would you choose now?

(A) *Problem I* (B) *Problem II*

Comparison IV

Instructions: Please first examine Comparison Problems **I, II, III and IV** below, and then answer questions 46-69 about them.

I. The weight of H_2SO_4 (molecular weight 98.1) in 50.0 milliliters of a 6.00-molar solution is

(A) 3.10 grams (B) 12.0 grams
(C) 29.4 grams (D) 294 grams
(E) 300. grams

II. If a copper sample containing some zinc impurity is to be purified by electrolysis, the anode and the cathode must be which of the following?

Anode	Cathode
(A) Pure copper	Pure zinc
(B) Pure zinc	Pure copper
(C) Pure copper	Impure copper sample
(D) Impure copper sample	Pure copper
(E) Impure copper sample	Pure zinc

III. A 0.20-molar solution of a weak monoprotic acid, HA, has a pH of 3.00. The ionization constant of this acid is

(A) 5.0×10^{-7} (B) 2.0×10^{-7}
(C) 5.0×10^{-6} (D) 5.0×10^{-3}
(E) 2.0×10^{-3}

IV. For the substance represented in the diagram on the *left*, which of the phases is most dense and which is least dense at -15°C?

Most Dense	Least Dense
(A) Solid	Gas
(B) Solid	Liquid
(C) Liquid	Solid
(D) Liquid	Gas

(E) The diagram gives no information about densities.

46. To what extent have you studied a problem like Problem I?

(Never) A B C D E (Extensively)

47. To what extent have you studied a problem like Problem II?

(Never) A B C D E (Extensively)

48. To what extent have you studied a problem like Problem III?

(Never) A B C D E (Extensively)

49. To what extent have you studied a problem like Problem IV?

(Never) A B C D E (Extensively)

50. How similar are Problem I and II in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

51. Which one seems easier for you?

(A) *Problem I*
(B) *Problem II*

52. How similar are Problem I and III in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

53. Which one seems easier for you?

(A) *Problem I*
(B) *Problem III*

54. How similar are Problem I and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

55. Which one seems easier for you?

(A) *Problem I*
(B) *Problem IV*

56. How similar are Problem II and III in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

57. Which one seems easier for you?

(A) *Problem II*
(B) *Problem III*

58. How similar are Problem II and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

59. Which one seems easier for you?

(A) *Problem II* (B) *Problem IV*

60. How similar are Problem III and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

61. Which one seems easier for you?

(A) *Problem III*
(B) *Problem IV*

62. How similar are Problem III and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

63. Which one seems easier for you?

(A) *Problem III*
(B) *Problem IV*

64. If you were allowed to choose ***three*** questions to answer, which three questions would you like to choose?

(A)= *Problem I II III*
(B)= *Problem I II IV*
(C)= *Problem I III IV*
(D)= *Problem II III IV*

65. Choose the correct answer to Problem I:

(A) (B) (C) (D) (E)

66. Choose the correct answer to Problem II:

(A) (B) (C) (D) (E)

67. Choose the correct answer to Problem III:

(A) (B) (C) (D) (E)

68. Choose the correct answer to Problem IV:

(A) (B) (C) (D) (E)

69. If you would like to choose three questions *again* after you have tried to answer the four questions, which three questions would you like to choose now?

(A)= *Problem I II III*
(B)= *Problem I II IV*
(C)= *Problem I III IV*
(D)= *Problem II III IV*

Part D: AP Chemistry
Essay Problem Comparisons

Instructions: In this section, there are five essay-type problems. Please first read through all five essays, and then answer the questions 70-95 about them.

Essay I: CF_4 XeF_4 ClF_3

(a) Draw a Lewis electron-dot structure for each of the molecules above and identify the shape of each.
(b) Use the valence shell electron-pair repulsion (VSEPR) model to explain the geometry of each of these molecules.

Essay II: The melting points of the alkali metals decrease from Li to Cs. In contrast, the melting points of the halogens increase from F_2 to I_2.

(a) Using bonding principles, account for the decrease in the melting points of the alkali metals.
(b) Using bonding principles, account for the increase in the melting points of the halogens.
(c) What is the expected trend in the melting points of the compounds LiF, NaCl, KBr, and CsI ? Explain this trend using bonding principles.

Essay III: $C_2H_4(g) + H_2(g) \rightarrow C_2H_6(g)$ $\Delta H^0 = -137kJ$

Account for the following observations regarding the exothermic reaction represented by the equation above.

(a) An increase in the pressure of the reactants causes an increase in the reaction rate.
(b) A small increase in temperature causes a large increase in the reaction rate.
(c) The presence of metallic nickel causes an increase in reaction rate.
(d) The presence of powdered nickel causes a larger increase in reaction rate than does the presence of a single piece of nickel of the same mass.

Essay IV: The carbon isotope of mass 12 is stable. The carbon isotopes of mass 11 and

mass 14 are unstable. However, the type of radioactive decay is different for these two isotopes. Carbon-12 is not produced in either case.

(a) Identify a type of decay expected for carbon-11 and write the balanced nuclear reaction for that decay process.

(b) Identify the type of decay expected for carbon-14 and write the balanced nuclear reaction for that decay process.

(c) Gamma rays are observed during the radioactive decay of carbon-11. Why is it unnecessary to include the gamma rays in the radioactive decay equation of (a) ?

(d) Explain how the amount of carbon-14 in a piece of wood can be used to determine when the tree died.

Essay V: Consider three unlabeled bottles, each containing small pieces of one of the following metals.

- Magnesium - Sodium - Silver

The following reagents are used for identifying the metals.

- Pure water
- A solution of 1.0-molar HCl
- A solution of concentrated HNO_3

(a) Which metal can be easily identified because it is much softer than the other two? Describe a chemical test that distinguishes this metal from the other two, using only one of the reagents above. Write a balanced chemical equation for the reaction that occurs.

(b) One of the other two metals reacts readily with the HCl solution. Identify the metal and write the balanced chemical equation for the reaction that occurs when this metal is added to the HCl solution. Use the attached table of standard reduction potentials to account for the fact that this metal reacts with HCl while the other does not.

(c) The one remaining metal reacts with the concentrated HNO_3 solution. Write a balanced chemical equation for the reaction that occurs.

(d) The solution obtained in (c) is diluted and a few drops of 1 *M* HCl is added. Describe what would be observed. Write a balanced chemical equation for the reaction that occurs.

70. To what extent have you studied a problem like Essay I?

(Never) A B C D E (Extensively)

71. To what extent have you studied a problem like Essay II?

(Never) A B C D E (Extensively)

72. To what extent have you studied a problem like Essay III?

(Never) A B C D E (Extensively)

73. To what extent have you studied a problem like Essay IV?

(Never) A B C D E (Extensively)

74. To what extent have you studied a problem like Essay V?

(Never) A B C D E (Extensively)

75. How similar are Essay I and II in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

76. Which one seems easier for you?
(A) *Essay I* (B) *Essay II*

77. How similar are Essay I and III in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

78. Which one seems easier for you?

(A) *Essay I* (B) *Essay III*

79. How similar are Essay I and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

80. Which one seems easier for you?
(A) *Essay I* (B) *Essay IV*

81. How similar are Essay I and V in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

82. Which one seems easier for you?
(A) *Essay I* (B) *Essay V*

83. How similar are Essay II and III in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

84. Which one seems easier for you?
(A) *Essay II* (B) *Essay III*

85. How similar are Essay II and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

86. Which one seems easier for you?
(A) *Essay II* (B) *Essay IV*

87. How similar are Essay II and V in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

88. Which one seems easier for you?
(A) *Essay II* (B) *Essay V*

89. How similar are Essay III and IV in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

90. Which one seems easier for you?
(A) *Essay III* (B) *Essay IV*

91. How similar are Essay III and V in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

92. Which one seems easier for you?

(A) *Essay III* (B) *Essay V*

93. How similar are Essay IV and V in content?

(A) *Not at all similar*
(B) *Moderately similar*
(C) *Extremely similar*

94. Which one seems easier for you?
(A) *Essay IV* (B) *Essay V*

95. If you were allowed to choose ***three*** out of the five essays, which of the following combination would you choose?

(A)=*Essay I II III*
(B)=*Essay I II IV*
(C)=*Essay I II V*
(D)=*Essay I III IV*
(E)=*Essay I III V*
(F)=*Essay I IV V*
(G)=*Essay II III IV*
(H)=*Essay II III V*
(I)=*Essay II IV V*
(J)=*Essay III IV V*

The End